(12) **EUROPEAN PATENT APPLICATION**

(72) Inventors:
• **Schriel, James Michael**
  Nepean Ontario K2C 3H2 (CA)
• **Megarity, Mark R.**
  Dunrobin Ontario K0A 1T0 (CA)

(74) Representative: **Feray, Valérie et al
Feray Lenne Conseil
44/52, Rue de la Justice
75020 Paris (FR)**

(54) **Method and apparatus capable of transferring very high data rates across a midplane or backplane**

(57) A method and apparatus of communicating data packets across the midplane of an electronic system in which the packets are partitioned into segments of a predetermined size and then serialized to a predetermined width. The serialized packets are transmitted, in phase staggered segments, across the midplane on a respective channel, received into receiving end and the serialized segments that have traversed the midplane, are deserialized and reassembled into the original data packet.

EP 1 424 871 A2

Description

[0001] This invention relates to a bus capable of transferring variable length packets (e.g. for POS) at a 10Gbps rate between two separate cards across a midplane or backplane.

[0002] The midplane or backplane (midplane/backplane) has a limited number of physical pins available and signals must pass through two connectors, which potentially could introduce signal integrity issues for high-speed signals.

[0003] There is a width vs. speed tradeoff, wherein a slower bus rate is easier and more reliable to implement however it must also be wider, which can be inefficient for short packets.

[0004] SQULB (prior art) is a sequenced utopia-3 like bus designed for asynchronous transfer mode (ATM) applications, i.e. fixed sized (56-byte) cells, and as such is not capable of handling variable length packets.

[0005] In order to modify SQULB to handle variable length packets, the bus must be made four times wider (four-bytes to sixteen-bytes). This solution is not feasible for the following reasons:

> 1. The limited number of pins available across the midplane/backplane. There is currently no offering of serializer/deserializer (SERDES) devices capable of handling this bus width. Separate SERDES devices would make it difficult to receive the packet data in the proper order and with minimal signal skew.
> 2. A sixteen-byte wide bus implies that a single packet could contain up to fifteen empty bytes transferred during the end of packet. This transmission model would be very inefficient and would require a large increase in the operating frequency of the physical bus to maintain a 10Gbps rate.

[0006] The present invention provides a method of communicating data packets across a midplane or a backplane of an electronic system, solving the problem that will be encountered by any efforts to pass variable-length packet data between separate cards.

[0007] More precisely, the present invention provides a method of communicating data packets across a midplane or a backplane of an electronic system, comprising the steps of:

> a) forming a plurality of channels through said midplane or backplane,
> b) partitioning each data packet into segments of a predetermined size;
> c) serializing each segment to a predetermined width;
> d) transmitting each serialized segment across said midplane or backplane on a respective channel,
> e) receiving, from their respective channels, serialized segments that have traversed the midplane or

backplane;
> f) deserializing the received serialized segments; and
> g) re-assembling the deserialized segments into the data packet.

[0008] The present invention is directed to a method for partitioning packets into segments of a predetermined size (e.g. 64 bytes), serializing the segments, and transmitting the segments over a plurality of channels (for example, four) that have a staggered phase relationship to one another, and wherein the phase difference between adjacent channels (actually, adjacent segments in the sequence of segments that form the packet) is more than the maximum latency that can occur in any one channel, thereby maintaining the ordering of the serialized packet segments. Preferably, there are 64-byte segments divided into four channels and staggering is 2 times maximum latency. While the invention is particularly applicable to variable length packets, the invention can also be used for fixed-length cells as well as variable length packets.

[0009] The advantages of the invention, in general, and over the closest prior art solution include:

> 1. Variable length packets can be transmitted very efficiently at a low bus frequency while maintaining a bandwidth of 10Gbps. Furthermore, in the present implementation, the segmentation of packets into 64-byte segments means that ATM cells can also be transmitted very efficiently over the bus.
> 2. The described bus topology is easily amenable to quad OC48.
> 3. Since the bus is divided into segments transmitted four-bytes at a time, there is only a need for a maximum of three empty bytes per packet. This scheme makes the bus more efficient and allows it to operate at a slower frequency while still achieving 10Gbps rates.

[0010] The above and other objects, advantages and features of the invention will become more apparent when considered with the following specification and accompanying drawings wherein:

> Figure 1 is a block diagram illustrating the packets divided into 64-byte segments for transmitting over four channels are staggered in their phase relationship to one another,
> Figure 2 is a timing chart showing a single channel bus timing of a given example, and
> Figure 3 illustrates the staggering performed by the transmitting device.

[0011] Referring to Figure 1, packets from transmitter T are divided into 64-byte segments for transmission over four channels CHA, CHB, CHC and CHD that are staggered in their phase relationship to one another. A

first array or set of SERDES (serializer/deserializer) devices TSA, TSB, TSC and TSD at the transmitter T in each channel is used to serialize the data and control for transmission across the midplane/backplane P and a second array or set of SERDES devices RSA, RSB, RSC and RSD are used at the receiver R in each channel to deserialize the data. These SERDES devices allow data and control to be passed across the midplane/backplane in a compressed manner and thus reduce the large variable amount of latency for a complete packet transfer. Furthermore, by staggering the phase relationship of the channels, the maximum latency that can occur in any one channel is accommodated and therefore the ordering of packet segments is maintained. For example, in the embodiment implemented in Figure 1, the channels are staggered by 40 ns and the maximum latency per channel is 20 ns.

[0012] The invention has two basic aspects: the topology of Figure 1 and a bus protocol that runs on that topology to provide the required bandwidth of 10Gbps for variable length packets.

[0013] With reference to Figure 2, the bus protocol will now be briefly described:

Egress Data Path Signal Definitions

[0014] The following describes a single channel and what is described is applicable for all four channels and can easily be extended to the complete bus.

[0015] Each channel includes a plurality of core data path signals, a 32-bit wide data bus with eleven-bits of out-of-band control, and a number of non-core data path signals, which may be used to transfer additional information.

[0016] Packets are broken apart into segments, 64-bytes of data, and transmitted 32-bits per cycle over 16 clock cycles. Packets that are greater than 64 bytes are required to be transmitted over more than one channel.

[0017] When data is available for transmission, a Start of Segment (SOS) is raised for one clock cycle concurrent with the first word of the transfer.

[0018] When a new packet is being transmitted, a Start of Packet (SOP) signal is raised for one clock cycle concurrent with the first word of the packet. Similarly, when the last word of a packet is being transmitted, an End of a Packet (EOP) signal is raised for one clock cycle concurrently with the last word. The Empty (MPTY) signals indicate how many bytes of the current word are valid. Since the bus is a word-wide (i.e. 32-bits will be transferred each clock cycle) up to three bytes of PAD may be present on a transfer. The MPTY signals are only valid when an EOP occurs. If the packet happens contain an error, then the Error (ERR) signal becomes active while EOP is active.

[0019] During the complete data transfer, the Valid (VAL) signal is active. The parity across this interface should always be valid for both the data path (DPRTY) signal and the control path (CPRTY) signal.

[0020] If there is no data to be transferred, idle cycles will occur on the bus. For an idle cycle, the data bus and all the control signals (except parity) will be driven low. The parity will remain valid at all times.

Summarizing:

[0021]

1. Data packets are divided into 64-byte segments,
2. This interface has parity protection for all data and control signals (DPRTY and CPRTY),
3. Each segment is serialized to four bytes wide [Data(31:0)] and each segment is tagged with a set of out-of-band control signals, as shown in Figure 2. The associated control signals consist of start and end of packet indications (SOP and EOP), error indication (ERR), and the number of empty bytes transmitted during the end of packet condition. This information is then used by the receiver to properly assemble the packet.

[0022] Figure 3 illustrates the staggering performed by the transmitting device T.

[0023] The advantages of the invention, in general, and over the closest prior art solution include:

1. Variable length packets can be transmitted very efficiently at a low bus frequency while maintaining a bandwidth of 10Gbps. Furthermore, in the present implementation, the segmentation of packets into 64-byte segments means that ATM cells can also be transmitted very efficiently over the bus.
2. The described bus topology is easily amenable to quad OC48.
3. Since the bus is divided into segments transmitted four-bytes at a time, there is only a need for a maximum of three empty bytes per packet. This scheme makes the bus more efficient and allows it to operate at a slower frequency while still achieving 10Gbps rates.

[0024] The invention is directed to the method and apparatus of partitioning packets into segments of predetermined size (for example, 64-bytes), serializing the segments, and transmitting them over a plurality of channels (for example, four) that have a staggered phase relationship to one another, and wherein phase difference between adjacent channels (actually, adjacent segments in a sequence of segments that form the packet) and more than the maximum latency that can occur in any one channel, thereby maintaining the order of the serialized packet segments. The invention is not limited to variable-length packets as the invention can be used for both fixed-length cells and variable-length packets.

[0025] This invention addresses an obstacle and

solves the problem that will be encountered by any efforts to pass variable-length packet data between separate cards.

[0026] While the invention has been described in relation to preferred embodiments of the invention, it will be appreciated that other embodiments, adaptations and modifications of the invention will be apparent to those skilled in the art.

Claims

1. Method of communicating data packets across a midplane or a backplane of an electronic system, comprising the steps of:

   a) forming a plurality of channels through said midplane or backplane,
   b) partitioning each data packet into segments of a predetermined size;
   c) serializing each segment to a predetermined width;
   d) transmitting each serialized segment across said midplane or backplane on a respective channel,
   e) receiving, from their respective channels, serialized segments that have traversed the midplane or backplane;
   f) deserializing the received serialized segments; and
   g) re-assembling the deserialized segments into the data packet.

2. Method defined in claim 1 wherein said step d) includes the step of delaying the transmission of a segment in one channel from that of a channel carrying a next sequential segment of the data packet so that the segments are staggered relative to each other.

3. Method defined in claim 1 or claim 2 wherein step d) further includes tagging the packet with the out-of-band control and error correction/detection signals.

4. Apparatus in a communication system having a digital data transmitter side apparatus for transmitting digital data packets of variable length between two separate points across a midplane or a backplane to a digital data receiver side apparatus, said midplane or backplane having a limited number of physical pins available, said apparatus comprising:

   means at said digital data transmitter for partitioning each digital data packet into segments of a predetermined size,
   an array of serializers connected to said transmitter side for serializing each segment into a predetermined width, each serializer being connected to a separate pathway through said midplane or backplane,
   an array of deserializers on said receiver side corresponding in number to the number of serializers on said transmitter side of said midplane or backplane and corresponding thereto, respectively, for receiving from their respective channels the serialized segments of data traversing said midplane or backplane and means for reassembling the deserialized segments into the data packet.

5. Apparatus defined in claim 4 wherein in said partitioning of packets into a predetermined size and serializing of segments thereof and transmitting said segments over a plurality of channels, said channels have a staggered phase relationship to one another wherein the phase difference between adjacent channels is more than the maximum latency that can occur in any one channel thereby maintaining the ordering of the serialized packet segments.

6. Apparatus defined in claim 5 wherein said data segments are in 64-byte segments and the number of channels in said midplane are four and staggering is two times the maximum latency.

7. Apparatus defined in any one of claims 4 to 6 wherein said transmission of segments in the respective channels are delayed from that of a channel carrying a next sequential segment of the packet.

8. Apparatus defined in any one of claims 4 to 7 including means for tagging the packets with an out-of-band control and error correction/detection signals.

9. Communication system having a digital data transmitter side apparatus for transmitting digital data packets of variable length between two separate cards across a midplane or a backplane to a digital data receiver side apparatus, said transmitter side apparatus including means for partitioning each digital data packet into segments of a predetermined size and an array of serializers connected to said transmitter side for serializing each segment into a predetermined width, each serializer being connected to a separate pathway through said midplane or backplane.

10. Receiver side apparatus in a communication system having a digital data transmitter side for transmitting digital data packets of variable length between two separate cards across a midplane or a backplane to a digital data receiver side apparatus and said transmitter side apparatus forming multiple channels,

said receiver side apparatus being comprised of an array of deserializers corresponding in number to the number of serializers on said transmitter side for receiving from their respective channels the serialized segments of data that traverse said midplane and reassembling the deserialized segments of data into data packets.
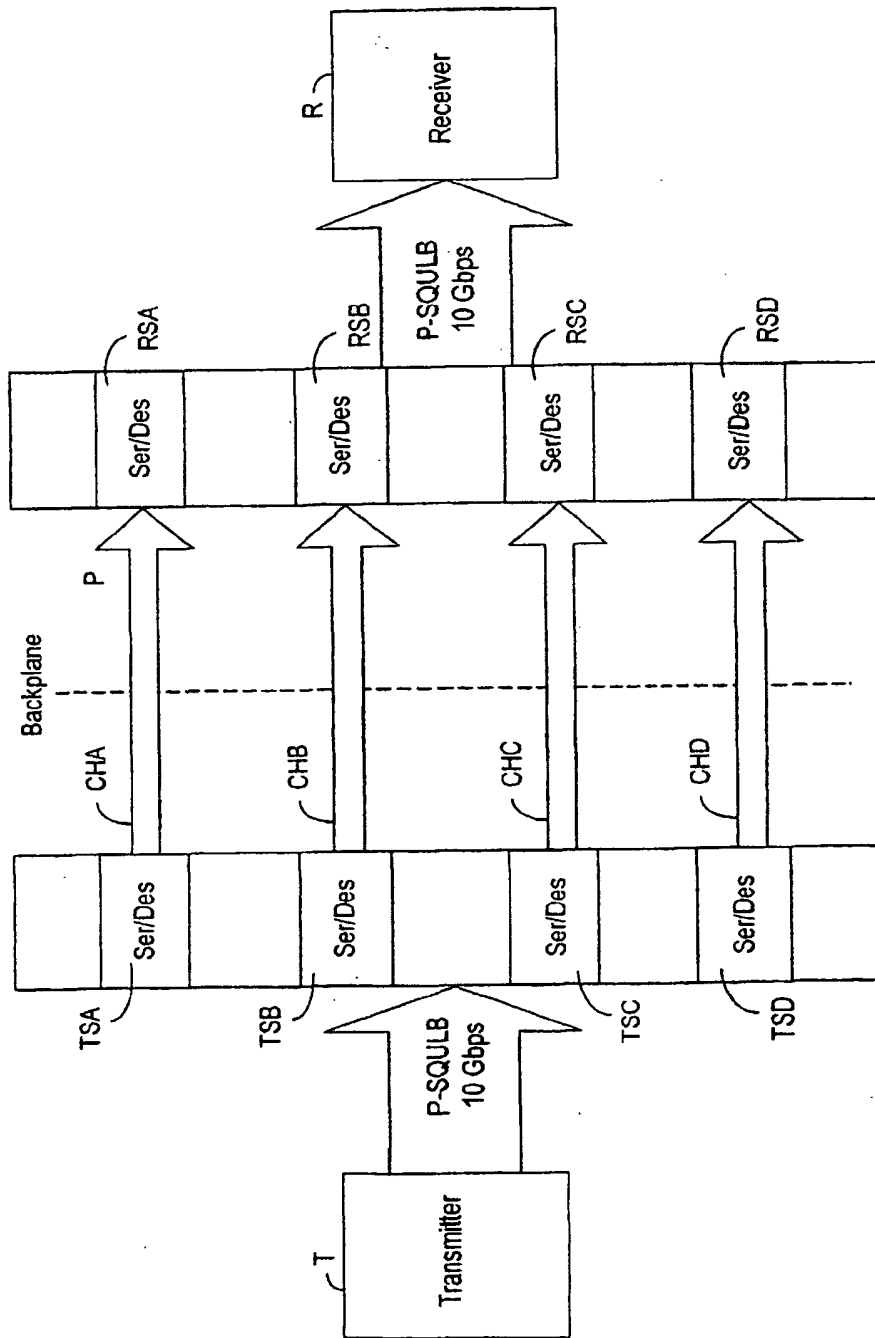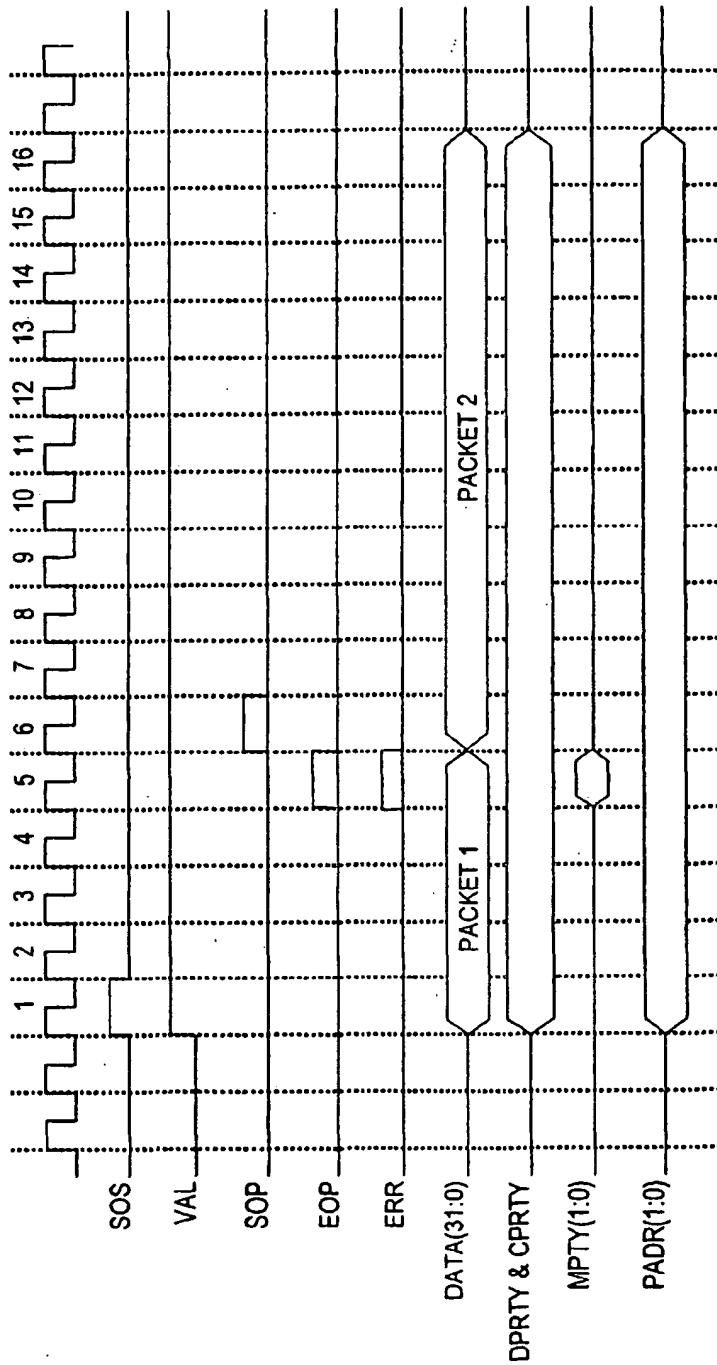
5

10

15

20

25

30

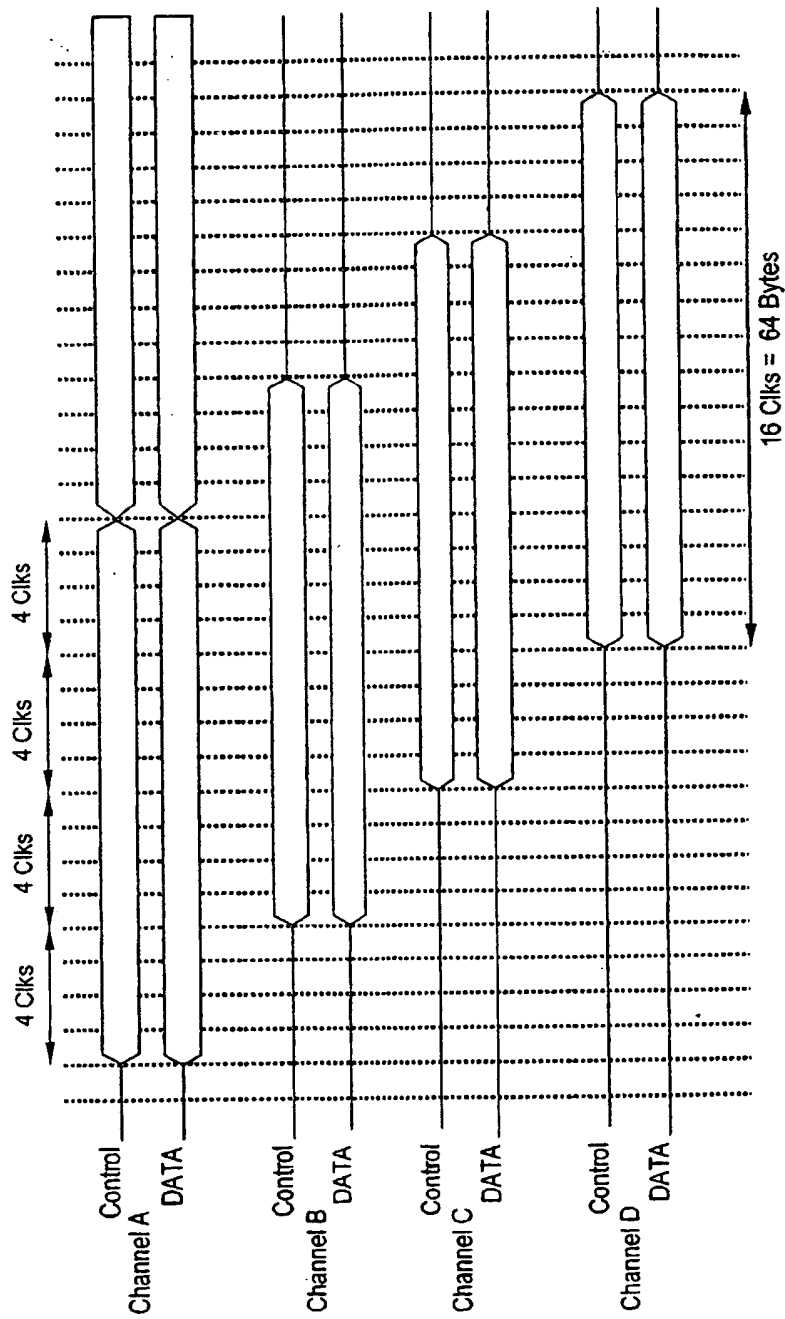35

40

45

50

55

FIG. 1

FIG. 2

SOS
VAL
SOP
EOP
ERR
DATA(31:0)
DPRTY & CPRTY
MPTY(1:0)
PADR(1:0)

PACKET 1
PACKET 2

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16

FIG. 3